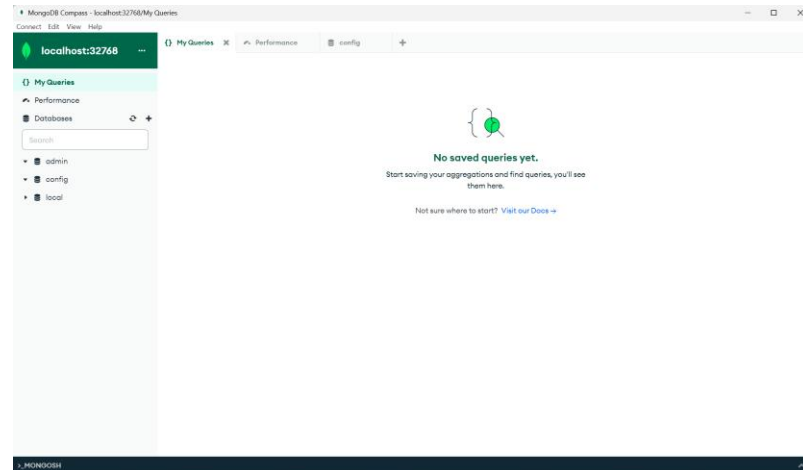


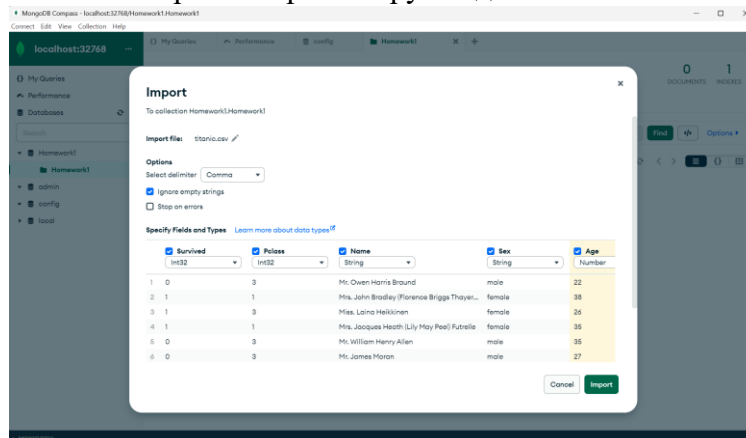
# Домашняя работа №1. MongoDB

## 1. Создание базы данных

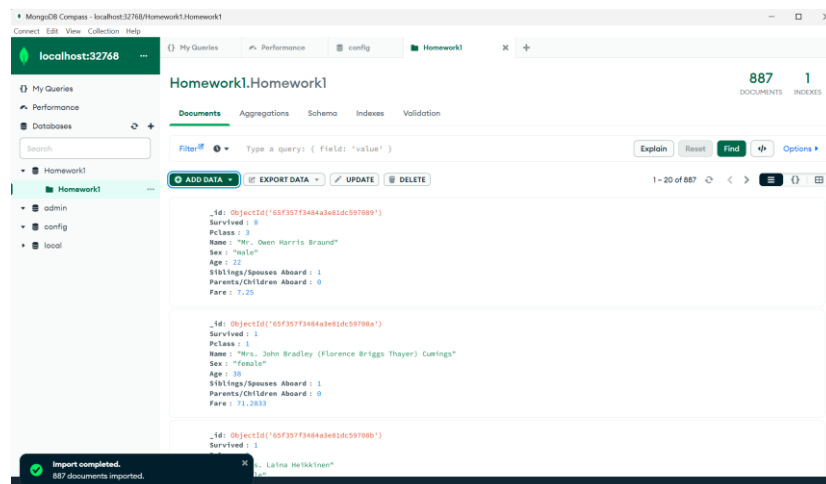
Сначала я начал работать с базой данных через MongoCompass. Нашел датасет и загрузил его через интерфейс.



После этого я также через Compass загрузил данные из тестовой выборки titanic.



Данные представлены в таком виде.



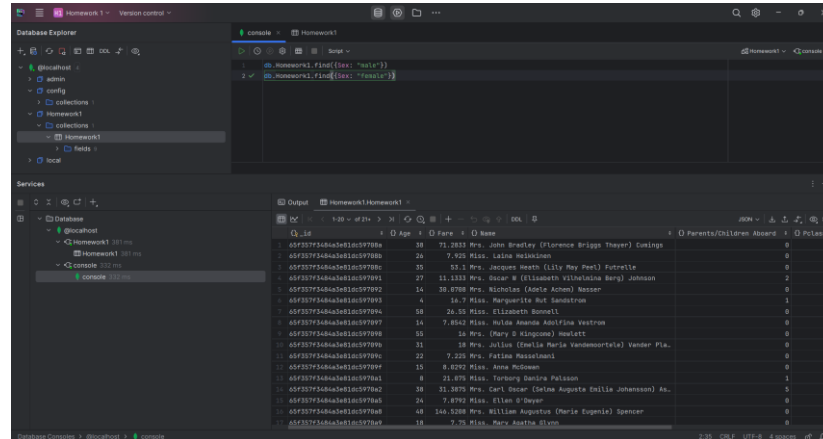
## 2. Выборка, агрегация, удаление и добавление данных

Здесь я решил, что удобнее будет пользоваться DataGrip.

Первый запрос выполняет простой поиск ячеек с определенным полем.

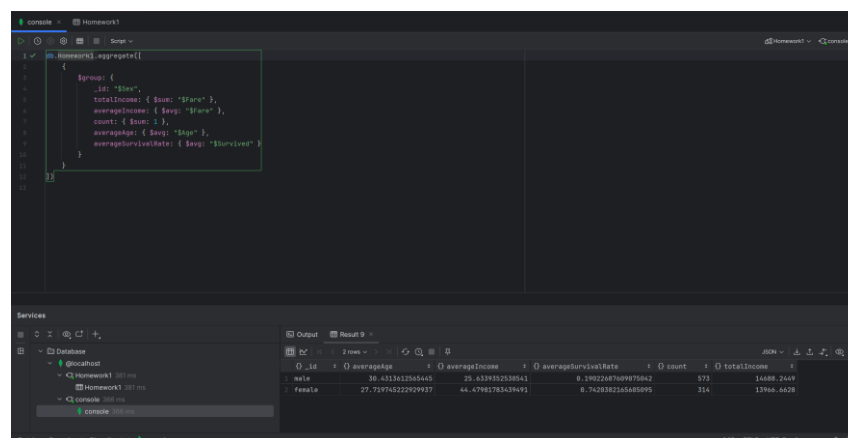
```
db.Homework1.find({ Genre: "Male" })
```

```
db.Homework1.find({ Genre: "Female" })
```



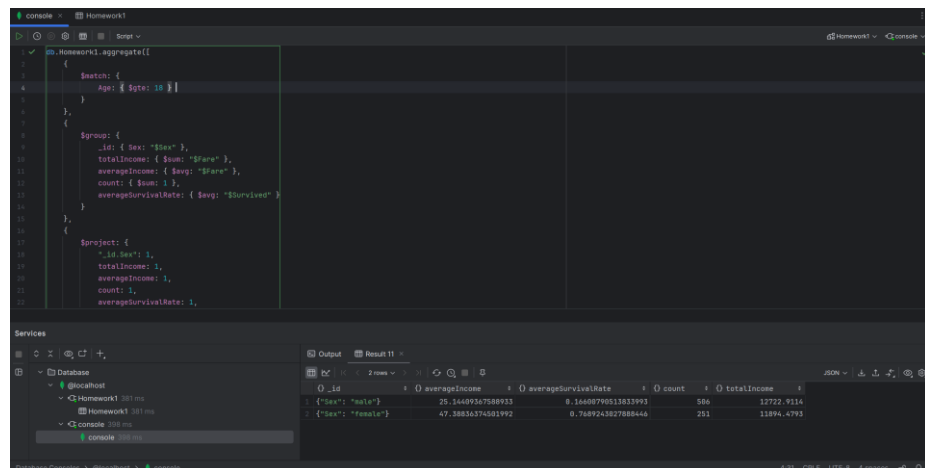
Сделаем запрос который агрегирует данные о пассажирах по полу и вычисляет различные статистические показатели для каждой группы пассажиров.

```
db.passengers.aggregate([
{
$group: {
_id: "$Sex",
totalIncome: { $sum: "$Fare" },
averageIncome: { $avg: "$Fare" },
count: { $sum: 1 },
averageAge: { $avg: "$Age" },
averageSurvivalRate: { $avg: "$Survived" }
}
}
])
```



## Чуть усложним запрос

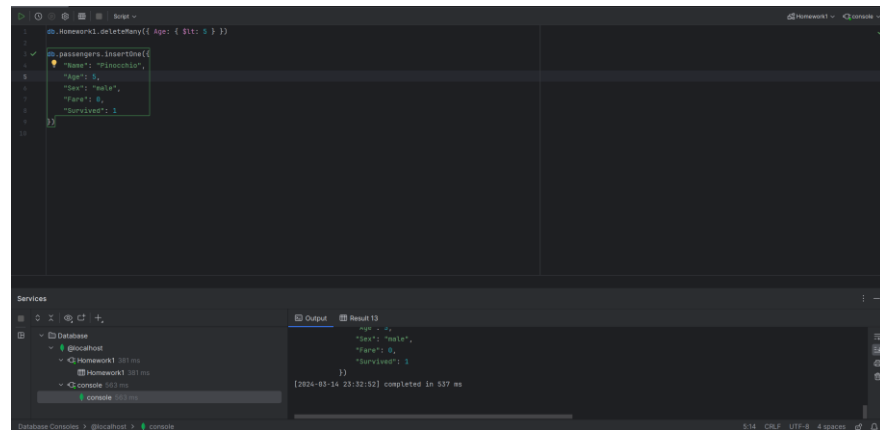
```
db.passengers.aggregate([
{
$match: {
Age: { $gte: 18 } // Фильтрация по возрасту
},
},
{
$group: {
_id: { Sex: "$Sex" },
totalIncome: { $sum: "$Fare" },
averageIncome: { $avg: "$Fare" },
count: { $sum: 1 },
averageSurvivalRate: { $avg: "$Survived" }
},
},
{
$project: {
"_id.Sex": 1,
totalIncome: 1,
averageIncome: 1,
count: 1,
averageSurvivalRate: 1,
}
}
])
```



Попробуем удалить всех пассажиров старше 5 лет, после добавим настоящего мальчика.

```
db.passengers.deleteMany({ Age: { $gt: 5 } })
db.passengers.insertOne({
  "Name": "Pinocchio",
  "Age": 5,
  "Sex": "male",
  "Fare": 0,
  "Survived": 1
})
```

}}



### 3. Индексы, разница производительностей

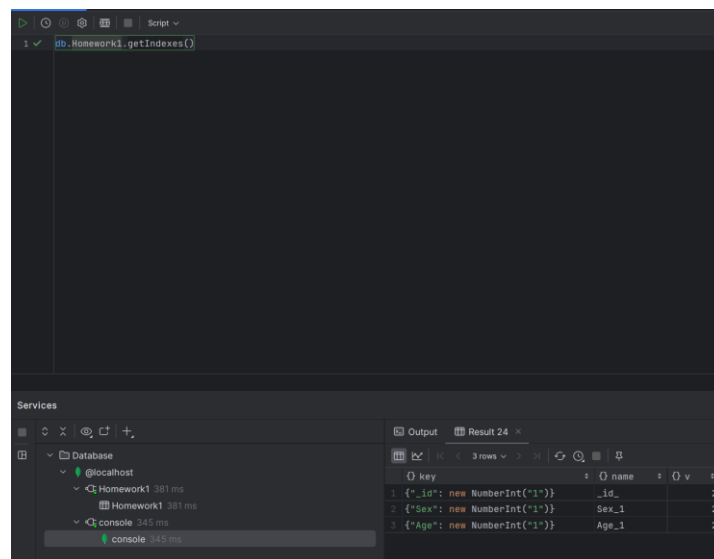
Добавим индексы

```
db.Homerwork1.createIndex({ Age: 1 })
```

```
db.Homerwork1.createIndex({ Sex: 1 })
```

Сделаем проверку индексов.

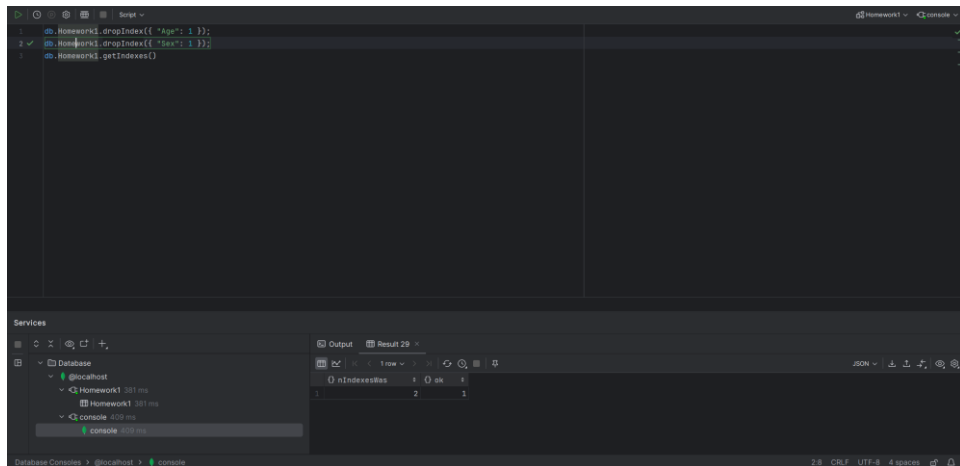
```
db.Homerwork1.getIndexes()
```



Удалим только что созданные индексы

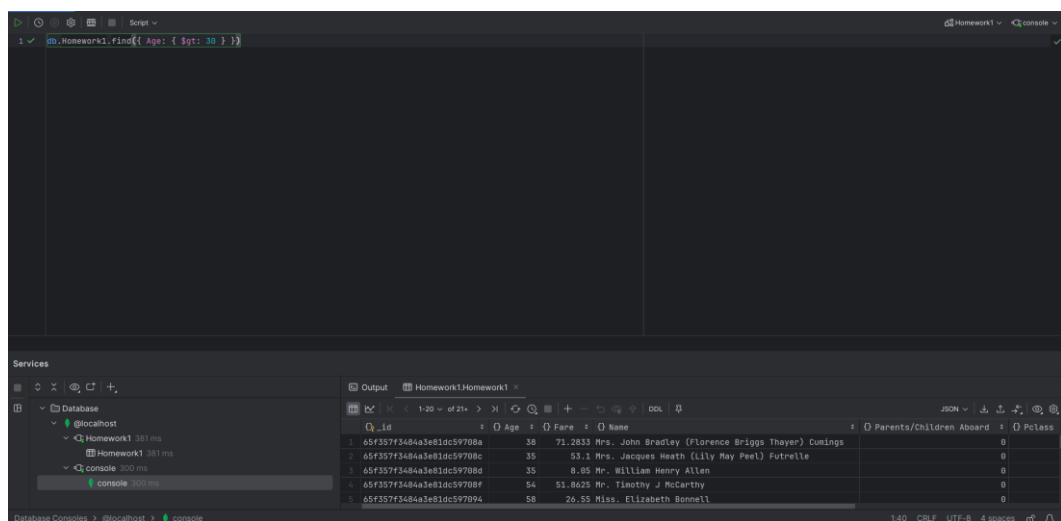
```
db.Homerwork1.dropIndex({ Age: 1 })
```

```
db.Homerwork1.dropIndex({ Sex: 1 })
```



Попробуем провести поиск.

`db.Homework1.find({ Age: { $gt: 30 } })`



После выполнения запроса время его выполнения составило приблизительно 300 мс. Перезапустим запрос с созданием новых индексов.

Несмотря на то, что индексы немного ускоряют работу, у нас наблюдается лишь небольшое улучшение. Это происходит из-за относительно небольшого объема данных. Важно также помнить, что при работе с большими объемами данных индексы могут создавать дополнительные нагрузки на память.

